input into the trigger input of the 555. The trigger input will produce pulses at the output of the 555 with the same frequency as the input signal. The pulses, however, will have a duration determined by the time constant RC. Thus, as the input frequency changes, the duty cycle of the 555 output will change. The FET acts as a constant-current source, with the 10 kΩ resistor providing the load. The voltage across the 10 kΩ resistor will be averaged by the 100 μF capacitor and the voltage at the output of the circuit will be determined by the duty cycle of the wave produced by the 555. Thus, the output is a voltage proportional to the input frequency. The values of R and C must be chosen to be compatible with the range of frequencies of interest. Several values of either R or C, or both, can be chosen using a switch.

4 Signal processing

4.1 Noise

Noise in the general sense refers to anything that obscures the quantity we are trying to measure. It can refer to interference, often from electromagnetic fields, or to random noise that is often thermal in origin. The various types of noise are discussed here.

a. Interference

In dealing with noise it is important to know the origin of the noise and its frequency spectrum. It is difficult to generalize in the case of interference. The nature of the interference depends on what we are trying to measure. Any outside parameter that couples to the parameter of interest is a concern. For example, if we were making sensitive Hall-effect measurements we might be concerned about the direction of the Earth's magnetic field, but we probably would not worry about a small amount of mechanical vibration from automobiles outside the building. On the other hand, if our aim were to perform optical interferometry we would not care about small magnetic fields, but too many cars passing by could be disastrous. The method of dealing with interference depends not only on the origin of the offending signal but on the nature of the experiment as well.

Sometimes the experiment can be isolated from the interference, as is done by placing optics experiments on massive stone tables or by placing r.f.-sensitive experiments inside a shielded (wire-mesh) room. If isolation is insufficient or inappropriate, then filtering sometimes works. This is particularly effective against 60 Hz interference from power cables. Since the frequency spectrum consists of a sharp peak at 60 Hz (and often additional sharp peaks at the harmonics of 60 Hz), a slot filter (band rejection) is often ideal.

Normal (passive) filters consist of combinations of resistors, capacitors and inductors arranged in such a way that the impedance does not permit signals of the undesired frequency to pass. Improved filtering characteristics can be obtained with active filters. These additionally use an operational effective to see whether we can eliminate its source. Sometimes, improperly placed power cables are responsible and we can easily separate signal-carrying cables from those that which are causing the interference. Ground loops are also often a problem. These are caused by connecting grounds in a circuit to different points. It is important in a circuit to have only *one* ground and to connect all ground leads to that one point.

Many of the techniques described later for noise reduction are suitable for eliminating interference.

b. *Johnson noise*

Johnson noise is the thermally induced noise that is present in any resistive element. It is the electronic analogue of the random motion of gas molecules in a box, which is known as Brownian motion. In noise terminology, Johnson noise is *white noise*. This means that the noise power per unit frequency is independent of frequency; that is, it has a *flat* frequency spectrum. To put it another way: if we were to measure the noise power between frequencies of, say, 100 Hz and 101 Hz, it would be the same as between, say, 1000 and 1001 Hz. The r.m.s. (root-mean-square) noise voltage is proportional to the square root of the resistance of the resistor, to the temperature and to the range of frequencies. We write

$$V_J = [4k_B TRB]^{1/2} \tag{4.1}$$

where T is in K, R in Ω, and B is the bandwidth in Hz. For an example, let us take a 10 KΩ resistor at room temperature. Over a band width of, say, 1 kHz the r.m.s. Johnson noise voltage is

$$V_J = 0.41\ \mu\text{V} \tag{4.2}$$

The 1 kHz band width can be in any part in the frequency spectrum; so we would measure 0.41 μV between 1 kHz and 2 kHz and also between 100 kHz and 101 kHz. This noise is present in all resistive elements and represents noise we could never eliminate. The Johnson noise is independent of the design of the resistor; so a cheap carbon resistor would have the same Johnson noise as an expensive precision wire-wound resistor of the same resistance. This is not necessarily true of other types of noise.

Although the r.m.s. voltage of the noise is given by equation (4.1), the distribution of noise voltages is a Gaussian centered at zero volts. That is, at any given time the probability that the noise will be in the voltage range V to $V + dV$ is given by

$$P(V, V + dV) = \frac{dV}{V_J\sqrt{2\pi}} \exp\left(\frac{-\frac{1}{2}V^2}{V_J^2}\right) \tag{4.3}$$

c. *Shot noise*

Shot noise is present in any circuit that carries a current. Its relative importance varies inversely with current. The origin of shot noise is as follows: A current I consists of a flow of electrons of the form

$$I = e\langle N\rangle \tag{4.4}$$

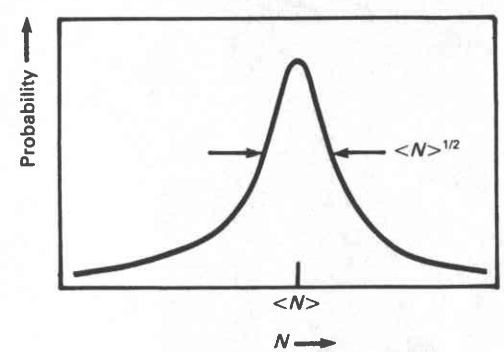
where $\langle N\rangle$ is the *average* number of electrons passing through a plane normal to the electron flow per unit time. For time in seconds and e in coulombs, I is in amperes. However, during any given time interval not necessarily exactly $\langle N\rangle$ electrons will pass the reference point. The distribution about $\langle N\rangle$ is given by $\langle N\rangle^{1/2}$, as for any statistical process; see Figure 4.1. This noise is a direct result of the quantization of the electric change and is given by (r.m.s. value)

$$I_s = (2eIB)^{1/2} \tag{4.5}$$

Like Johnson noise, shot noise is white noise. The shot noise as a percentage of the circuit current measured over, say, 10 kHz as a function of I is shown in Figure 4.2. We see that shot noise can be important for currents of less than one nanoampere or so and is a serious problem in the picoampere to femtoampere range.

d. *1/f noise*

$1/f$ noise is generally the most significant contribution to the noise (excepting interference). It depends a good deal on the construction of the components and the quality of the contacts. It can, unlike shot and Johnson noise, be reduced by careful choice of components. Its power spectrum is proportional to $1/f$ (hence the name). Given this dependence, the noise energy per frequency decade (rather than per unit frequency) is constant over the spectrum (over which the noise extends—it can be cut off below



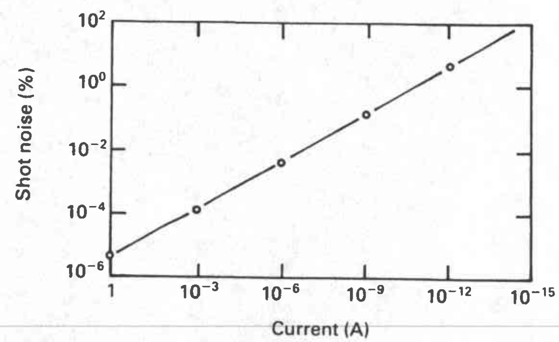


Figure 4.2. Relative importance of shot noise for different currents over a band width of 10 kHz.

and/or above certain frequencies). This type of noise spectrum is called *pink* noise.

For a resistor, the $1/f$ noise is proportional to the DC current flowing through it. The typical noise per frequency decade ranges from about 10^{-6} percent for high-quality wire-wound resistors to about 5×10^{-4} percent for inexpensive carbon composition resistors.

e. *Signal-to-noise ratio*

It is convenient to measure the quality of the signal in terms of a signal-to-noise ratio S/N defined in decibels (dB) to be

$$S/N = 10 \log_{10} \frac{V_s^2}{V_n^2} \tag{4.6}$$

where V_s and V_n are the signal and noise voltages, respectively, measured over the same bandwidth. If the signal has a frequency spectrum that extends over a limited bandwidth, then increasing the bandwidth beyond this point will only decrease the S/N. This is because, in general, V_n will increase without increasing V_s.

In the remainder of this chapter we discuss various methods for improving the S/N of an experiment.

4.2 Filters

There are all kinds of filters for all kinds of applications and there are elaborate and complex methods for designing the best filter for your needs. In view of this, it is not surprising that entire books (a lot of them) have been written on filters. Some examples are given in the references. We will

and then give some practical general-purpose filter circuits and discuss their characteristics.

We divide filters into two general categories: (1) passive filters and (2) active filters. Passive filters consist only of resistors, capacitors, and inductors; active filters consist of, among other things, operational amplifiers. We begin with passive filters.

a. *Passive filters*

An analysis of the characteristics of a filter circuit requires some knowledge of the response of circuit elements to AC signals. The general rules for the analysis of filter circuits are given in this section. The reader who is interested in a more thorough introduction to AC circuit theory can find this information in many introductory physics texts. The references for this chapter give three texts that are particularly recommended—Halliday and Resnick's *Physics*, Reimann's *Physics: Electricity, Magnetism and Optics*, and Purcell's *Electricity and Magnetism*. At a more advanced level, a good text is Kerchner and Corcoran's *Alternating-current Circuits*.

In a DC circuit the total resistance is the sum of the individual resistances in the circuit, summed in an appropriate way that depends on whether they are connected in series or parallel. In an AC circuit the impedance is analogous to the resistance and can be summed in a similar manner. It is important to note that the AC impedance has not only a real component but an imaginary component as well. It is this imaginary component that is responsible for the phase shifts in an AC circuit. The various components we are interested in are resistors, inductors, and capacitors. The contributions to the total impedance for those components are as follows.

1. The resistance

$$Z_R = R \tag{4.7}$$

2. The inductive reactance

$$Z_L = i\omega L \tag{4.8}$$

3. The capacitive reactance

$$Z_c = (i\omega c)^{-1} \tag{4.9}$$

It should be noted that for real components these three contributions cannot necessarily be separated. that is, a real inductor will have not only inductance but resistance and probably capacitance as well. Here we deal with ideal components.

Consider the simple circuit shown in Figure 4.3. This is a voltage divider circuit consisting of a resistor and a capacitor in series, the output taken across the capacitor. We can write the total impedance seen by V_{in} as

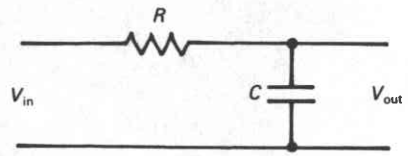


Figure 4.3. A simple first-order low-pass filter.

The ratio of V_{out} to V_{in} is given by

$$\frac{V_{out}}{V_{in}} = \frac{1}{i\omega C}\left(R + \frac{1}{i\omega C}\right)^{-1} \tag{4.11}$$

We can now solve for V_{out} in terms of V_{in}. In the present case we are primarily concerned with the amplitude of V_{out} and not its phase, so we write

$$|V_{out}| = (V_{out}V^*_{out})^{1/2} \tag{4.12}$$

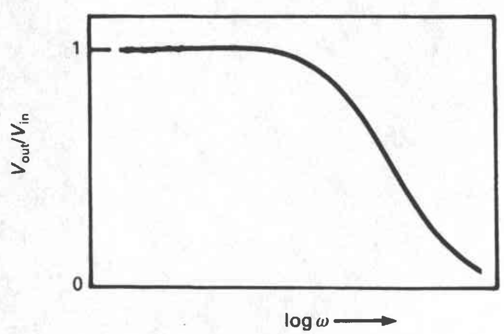
and from (4.11) we find

$$|V_{out}| = |V_{in}|\left[\frac{1}{\omega^2 C^2 R^2 + 1}\right]^{1/2} \tag{4.13}$$

We can now plot this function in terms of $\omega = 2\pi f$; we show this in Figure 4.4. As is customary, we have plotted this in terms of $\log \omega$. We see that $V_{out} = V_{in}$ for $\omega \ll 1/RC$; that is, all the voltage is passed at low frequencies. For $\omega \gg 1/RC$ all the voltage is dropped across the resistor.

For filters the ratio of input to output voltages is generally quoted in decibels, defined by equation (4.6). For the low-pass filter, equation (4.13) shows that at $\omega = 1/RC$,

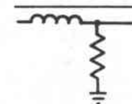
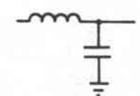
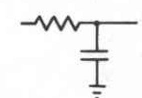
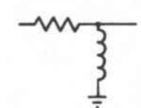
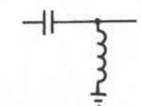
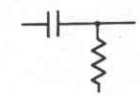
$$20 \log(2^{-1/2}) = -3 \text{ dB} \tag{4.14}$$



That is, the output is down by 3 dB at this frequency. The amount by which the output has decreased is called the *rolloff*. The rolloff is generally measured in dB per octave or dB per decade (octave = a factor of 2 in frequency; decade = a factor of 10 in frequency). For $\omega > 1/RC$ we find that for this filter the rolloff is 6 dB/octave or 20 dB/decade. This is typical of simple filters. These two component filters are referred to as first-order filters, and they come in low-pass and high-pass varieties. The various combinations that can be used and the frequency response of each are illustrated in Table 4.1. A band pass or a band reject filter can be constructed by using three (or more) components. Two simple examples are shown in Figures 4.5 and 4.6 along with their frequency responses.

Several first-order filters of the same type can be combined to produce a filter of a higher order: n first-order filters make an nth order filter. In general, the higher the order the sharper the rolloff, usually of 6 dB/octave per order. How sharply a particular filter needs to rolloff depends on how close the bandwidth of the signal you are interested in observing is to the noise you are interested in eliminating. For example, you do not need a

Table 4.1. First-order filter circuits

Circuit	Filter type	V_{out}/V_{in}
	Low pass	$\left(1 + \frac{\omega^2 L^2}{R^2}\right)^{-1/2}$
	Low pass	$\lvert(1 - \omega^2 LC)^{-1}\rvert$
	Low pass	$(1 + \omega^2 R^2 C^2)^{-1/2}$
	High pass	$\left(1 + \frac{R^2}{\omega^2 L^2}\right)^{-1/2}$
	High pass	$\left\lvert\left(1 - \frac{1}{\omega^2 LC}\right)^{-1}\right\rvert$
	High pass	$\left(1 + \frac{1}{\omega^2 R^2 C^2}\right)^{-1/2}$

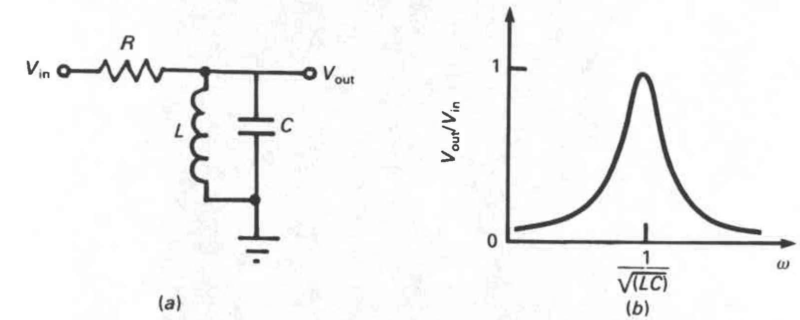
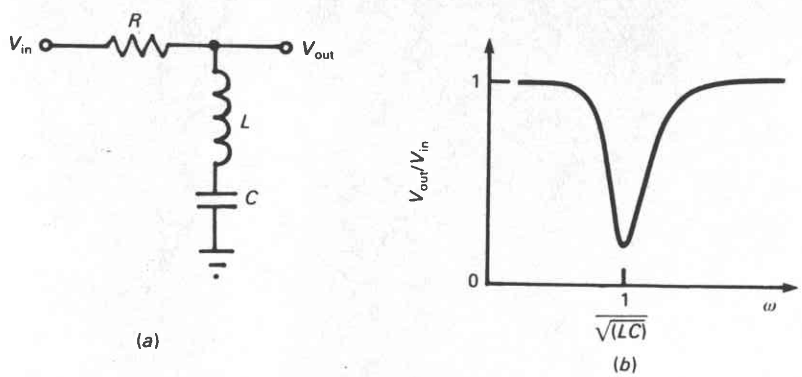


Figure 4.5. (*a*) A simple band pass filter; (*b*) its frequency response.

very good filter if you want to remove 60 Hz interference from a 10 kHz signal. On the other hand, if you want to extract a 100 Hz signal from 60 Hz interference then you need something much better.

There are numerous methods of putting together different combinations of first-order filters to form higher-order filters that have names such as *Butterworth filters, Chebychev filters,* or *Bessel filters*. Although active filters are generally better at fairly low frequencies, the rolloff of most op-amps at 100 kHz or so makes active filters useless above this. Thus, higher-order passive filters are best for higher frequencies. Similar higher-order band pass and band reject filters can be constructed as well.

Rather than going into extensive design criteria for filters in general we discuss the practical applications of one particular filter, the 60 Hz band reject filter. Figure 4.7 shows a circuit that is a good example of this kind of filter. Filter performance depends on the source and load impedances. We will consider a typical situation: a source impedance of 50 Ω and a load impedance (of an oscilloscope, for example) of 1 MΩ. Figure 4.8 shows the



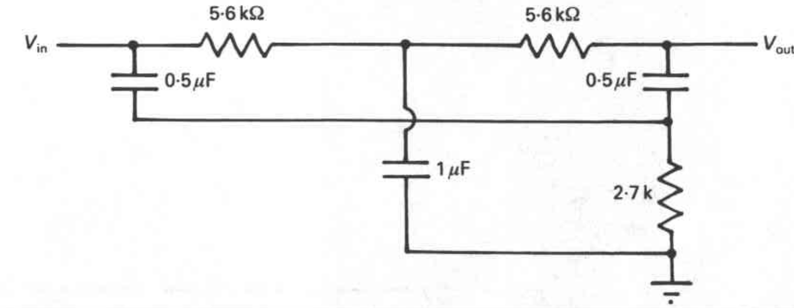


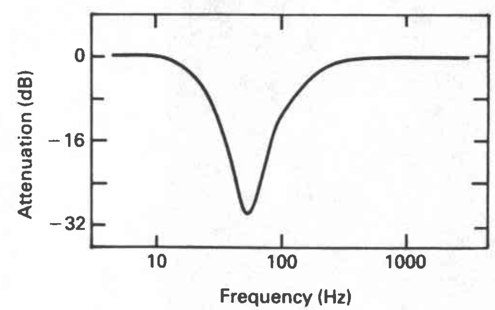
Figure 4.7. A 60 Hz band reject filter.

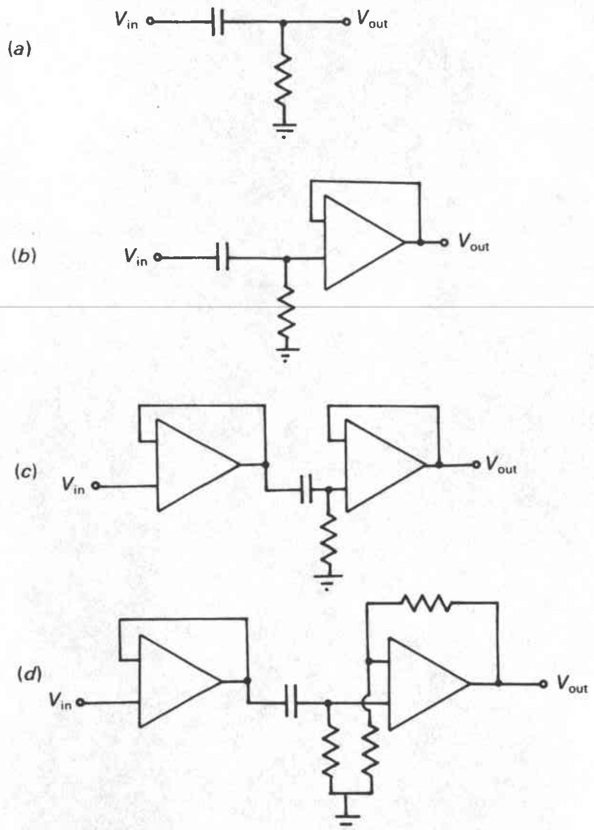
performance of this filter and demonstrates the usefulness of this type of circuit in rejecting noise at a particular frequency while allowing components of the signal at other frequencies to be passed.

b. *Active Filters*

There are a variety of ways of using op-amps to make better filters. Two problems that arise in the use of filters can be solved easily with op-amps: (1) the dependence of the filter characteristics on the input and output impedance; and (2) the attenuation of the signal due to the impedance of the filter. Let us see how op-amps help solve these problems.

In Figure 4.9*a* we show a conventional *RC* high-pass filter. In Figures 4.9*b–d* we show some analogous active filters. In the circuit in Figure 4.9*b* the op-amp is used to buffer the filter from the load. In this way the load impedance has no effect on the characteristics of the filter circuit. In this particular case the op-amp circuit is of unity gain. This op-amp configuration is called a *voltage follower*. Figure 4.9*c* shows a similar circuit in which the effects of source impedance have also been buffered by a voltage





igure 4.9. High-pass filters: (*a*) passive; (*b*) active with load buffering; (*c*) active with source nd load buffering; (*d*) active filter with additional gain.

'ollower. In Figure 4.9*d* we show a filter in which the output has been ;upplemented by an adjustable-gain op-amp. Figure 4.10 illustrates how ıigher-order filters can be constructed using op-amps.

There are a number of other ingenious things we can do with op-amps as :ar as filters are concerned. One of these is illustrated in Figure 4.11. There ıre source and load buffers at the input and output, respectively. At low frequencies the signal is blocked from entering the center op-amp by the :apacitor C_1; it therefore passes through R_3 to the output buffer. In this sense it is a conventional *RC* low-pass filter. Now comes the tricky part: high-frequency signals pass through C_1 and also through R_3. Through R_3 they go directly to the output buffer. From C_1 they are inverted by the

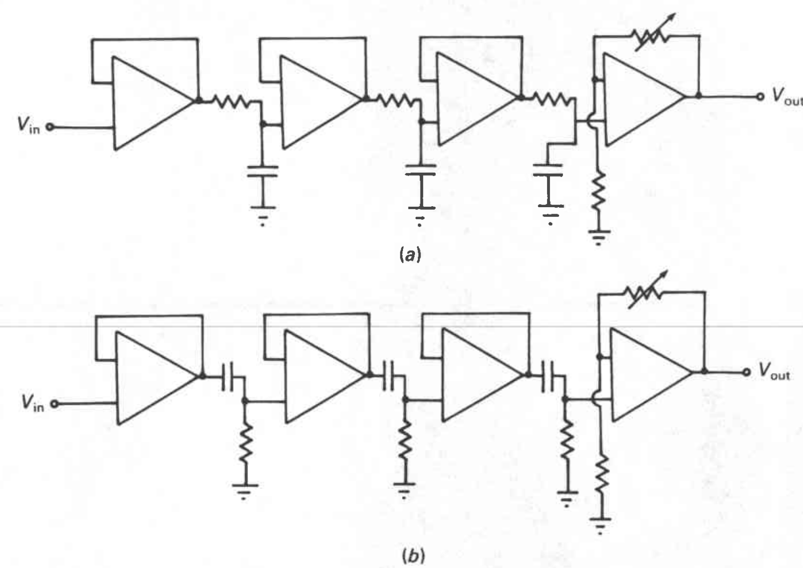


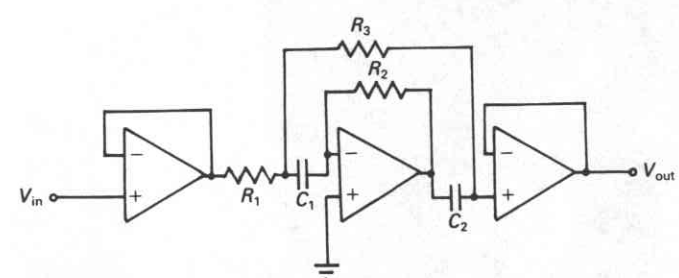
Figure 4.10. Third-order active filters: (*a*) low-pass; (*b*) high pass.

high-frequency signal is inverted it tends to cancel the high-frequency signal that arrived through R_3. This significantly increases the rolloff.

The filter design books listed in the references will serve as a guide for more circuits and to help you choose filter circuit components properly.

4.3 Signal averaging

Signal averaging is applicable to many types of experiments. Let us begin with some general consideration on how we go about making an experimental measurement.



In many cases a physics experiment does not consist of a single measurement of a particular quantity, but rather of a series of measurements of a quantity as a function of another quantity. For example we may want to measure the current through a semiconducting device as a function of the bias voltage.

The quantity we control in the experiment is referred to as the *extrinsic* parameter. The quantity we measure is referred to as the *intrinsic* parameter. In many experiments it is the case that there is more than one extrinsic parameter. It may be that there is more than one intrinsic parameter. Which quantities are intrinsic and which quantities are extrinsic is determined by the nature of the specific experiment. For example, current could be the intrinsic variable (and voltage the extrinsic variable) as in the example of the measurements on the semiconducting device given above. On the other hand, if we wanted to study the magnetic field H produced by a solenoid, the current might be the extrinsic variable and H the intrinsic one.

Conventionally, the extrinsic parameter x is varied over some relevant range of values and the intrinsic parameter y is measured. Signal averaging refers to the situation in which the extrinsic parameter is varied repeatedly over the same range of values and the intrinsic parameter is averaged (or summed) for each respective value of the extrinsic parameter. Mathematically we write

$$y = f(x_0) \tag{4.15}$$

The measured value of y in a single measurement is actually

$$y = f(x_0) + N \tag{4.16}$$

where N is the noise at the time the measurement was made. If we make n measurements of y at x_0 then we find

$$\langle y \rangle = \frac{1}{n}\sum f(x_0) + \frac{1}{n}\sum N \tag{4.17}$$

or

$$\langle y \rangle = f(x_0) + \langle N \rangle \tag{4.18}$$

The quantity $\langle N \rangle$ depends upon the nature of the noise in question, but for Poisson-distributed noise (i.e., random events), we find that

$$\langle N \rangle \propto n^{-1/2} \tag{4.19}$$

Therefore, as n increases, $\langle N \rangle$ decreases relative to $f(x_0)$. In fact from equation (4.6) we see that n measurements will improve the S/N relative to the S/N from a single measurement, $(S/N)_0$, by

$$S/N = 10 \log_{10} n(S/N)_0^2 \tag{4.20}$$

analyser. This is sometimes referred to as multichannel scaling. The *multichannel analyser* (*MCA*) consists of v channels or memory locations. Most commonly v is a power of 2; typically 256, 512, 1024, 2048 or 4096. Often the memory can be subdivided into 2, 4, 8 or even 16 subgroups, each of which can be used independently to store data. These channels are generally numbered from 0 to 255, 511, 1023, 2047 or 4095, respectively. The MCA begins in channel 0. During some specified time τ, known as the *dwell time*, the MCA will remain in channel 0 and will count pulses. These are added to the "sum" in the specific memory location associated with channel 0. The value of τ can be specified by the user within certain limits typically from ~0.1 ms to a few seconds. After remaining in the 0th channel for time τ, the analyser moves to channel 1. During the next τ all counts are added to the sum in the memory location associated with channel 1. This process repeats until channel $v - 1$ is reached. The entire process takes a time τv. At this this point the process repeats, beginning again in channel 0. Most MCAs have two modes: in one (*recurrent*) a sweep is initiated by the MCA immediately after the proceeding one has been completed; the other mode (*triggered*) requires an external signal to initiate each sweep. In some cases the first channel (channel 0) or even the first two or three channels are reserved for information (such as the number of sweeps completed) and are not used to accumulate data.

The object of the experiment is to produce an extrinsic parameter that is commensurate with the sweep of the MCA. That is, every time the MCA is in some channel i, the extrinsic variable should have the same value. This means that the intrinsic parameter (noise excluded) will have the same value every time the MCA is in channel i. Thus, performing each sweep is like making a series of measurements of y versus x, and while the real part of the signal will continue to add linearly in the number of sweeps the noise will add only according to equation (4.19) and the S/N will increase as equation (4.20).

Three aspects of the MCA need consideration (a) the nature of the input, (b) the method of synchronizing the MCA sweep to the extrinsic parameter and (c) the method of outputing the data from the MCA.

a. *Input*

Some experiments, such as many nuclear physics experiments, produce pulses directly; in many cases these can be input into the MCA directly. More commonly, in other areas of physics, the signal we are interested in measuring is an analog signal. This must be converted into a series of pulses in order to be recorded by the MCA. If we are interested in measuring a voltage, we want the number of pulses input to the MCA per unit time to be proportional to that voltage. The device most commonly used for this purpose is a V/F (*voltage-to-frequency converter*) or a VCO (*voltage-*

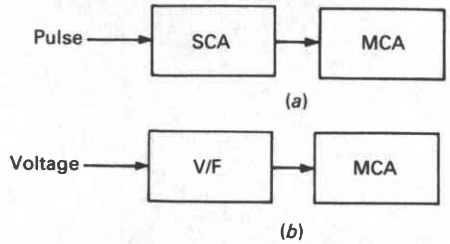


Figure 4.12. MCA input configurations for (*a*) pulse input and (*b*) voltage input.

are shown in Figure 4.12. In the pulse-detecting mode a single-channel analyser (SCA) is typically placed on the input. This enables us to set maximum and/or minimum pulse voltage that will be counted by the MCA. More details on the SCA and its operation will be given in the section on nuclear instrumentation (see Chapter 12).

b. *Synchronization*

There are various methods of controlling the extrinsic variable in a way that is synchronized with the sweep of the MCA. The best method is determined by the particular quantity we want to control as well as by the nature of the experiment. Most MCAs provide a various signals that can be used for this purpose. Figure 4.13 shows some examples for a hypothetical 16-channel MCA. (MCAs this small do not exist, but the picture would be too messy if we used 512 channels). Except for the ramp, most signals are TTL-compatible† (0–5 V). We give an example of the kind of things that can be done.

MCA example. We wish to vary an externally applied magnetic field from $-B$ to $+B$. We decide to ramp the field up and back down during each MCA sweep. To do this we use the circuit shown in Figure 4.14, which uses the MSB as the trigger. The first op-amp inverts the signal and shifts the voltage so the square wave is symmetric about zero. The second op-amp integrates the square wave to provide a triangular wave, and the third op-amp shifts the triangular wave so that it is symmetric about zero. Figure 4.15 shows the voltages at various points in the circuit. This output drives a programmable power supply (+ and − voltage output) for an electromagnet. The choice of the components in the integration circuit will depend on the dwell time and the characteristics of the programmable power supply. R' depends on R, C and τ. The voltage V_1 in Figure 4.15 is found to be

$$V_1 = 1.25\tau v/(RC) \tag{4.21}$$

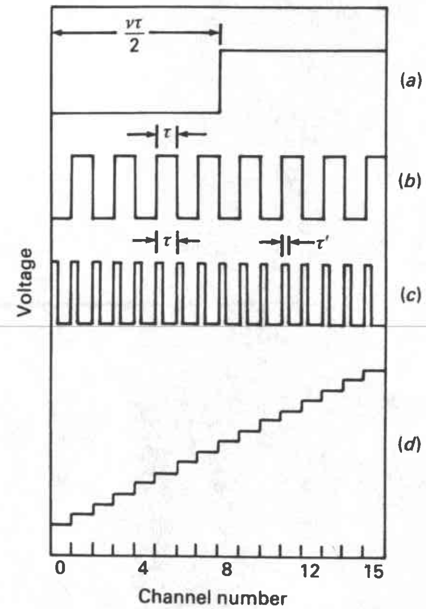


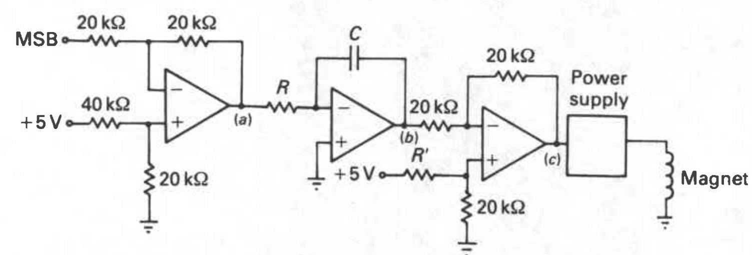
Figure 4.13. Typical MCA output signals for a hypothetical 16-channel analyser (τ = dwell time per channel): (*a*) most-significant bit (MSB); (*b*) least-significant bit (LSB); (*c*) channel advance signal; (*d*) digital ramp.

and from this R' is

$$R' = (5)(20\ \mathrm{k\Omega})/V_1 \tag{4.22}$$

c. *Output*

Essentially all modern MCAs have a CRT, such as an oscilloscope, to display the data stored in each of the channels. Many MCAs have an on-screen readout of one channel at a time. A cursor allows the channel of



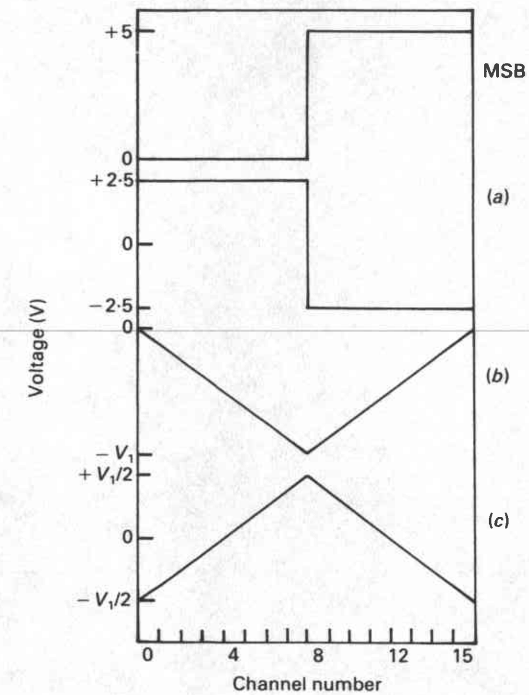


Figure 4.15. Voltages at the indicated points in the circuit of Figure 4.14.

interest to be selected. Most commonly, the data are output via a serial interface (RS-232) to a printer or to microcomputer or mainframe computer. The details of computer interfacing will be dealt with in Section 4.7. Some (more expensive) MCAs have a built-in interface for an x–y recorder. This allows an immediate hard copy of the data, plotted on an appropriate scale, to be obtained.

4.4 Pulse-height analysis

In nearly all cases, pulse-height analysers (PHAs) and multichannel analysers are incorporated in the same piece of equipment. Thus the term MCA is generally used to refer to a device that has these two modes of operation. You will usually see a switch somewhere on the device that enables you to change the mode.

The MCA accepts pulses in the PHA mode but the *voltage* of the pulses is the important factor. The v channels of the MCA represent different

with $V = V_{max}$, which corresponds to channel v. Each channel i represents a voltage range of V to $V + \Delta V$, where

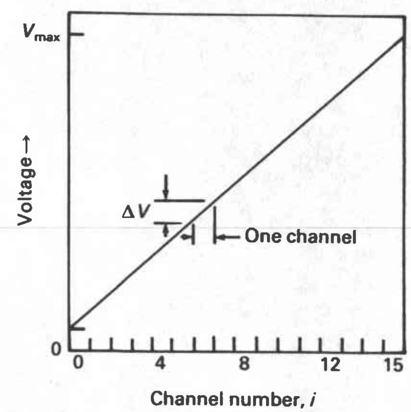
$$V = \frac{iV_{max}}{v} \tag{4.23}$$

and

$$\Delta V = \frac{V_{max}}{v} \tag{4.24}$$

The example of the hypothetical 16-channel MCA is illustrated in Figure 4.16. For each pulse that enters the MCA with a particular voltage V, one count is added to the sum in the memory location that corresponds to the channel representing a voltage range that includes V. V_{max} is chosen to include the voltage of the largest voltage pulses of interest. Most analysers have $V_{max} = 10$ V. In many cases this can be adjusted by adjustment of the conversion gain. If we cannot adjust V_{max} to a reasonable value with controls available on the MCA, then we can always modify the range of the voltage of the pulses by the insertion of an amplifier.

In the PHA mode the MCA will continue to accept and process pulses until the device is manually turned off or until some preset time limit is reached. It is common to reserve channel 0 for a number equal to the length of time in seconds for which the device has been accumulating data. The result is that we get a histogram of the number of pulses as a function of their voltage. As the number of available channels is large, ΔV is small, and very good voltage resolution in the histogram can be obtained. As in Figure 4.12 an SCA can be inserted before the analyser to block out pulses of too high or too low a voltage.



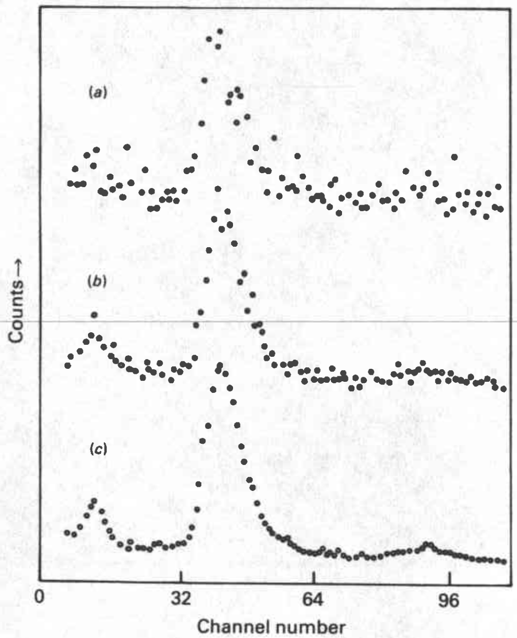


Figure 4.17. Example of noise elimination by signal averaging: spectrum accumulated for (*a*) 10 s; (*b*) 100 s, and (*c*) 1000 s.

Pulse height analysers are invaluable in the field of nuclear physics but is useful in many other fields as well. The energy spectra of radioactive sources, as shown in Chapter 12, are collected with a pulse-height analyser.

As with the MCA, the ability of the PHA mode to deal with noise relies on the statistical nature of the process producing the pulses and the fact that the signal will accumulate faster than the noise. Figure 4.17 illustrates this point. In Figure 4.17*a* the spectrum of the radioactive source was accumulated for 10 s; in Figure 4.17*b* accumulation proceeded for 100 s; in Figure 4.17*c* accumulation proceeded for 1000 s. Observe the clear increase in the S/N of the spectrum.

4.5 Phase-sensitive detection

Phase-sensitive detection is an effective method of eliminating noise from many types of experiments. Although it is particularly effective when used in conjunction with multichannel scaling, it is frequently useful in its own

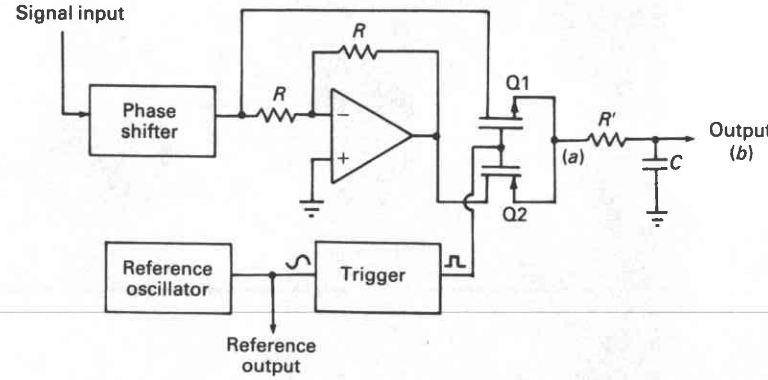


Figure 4.18. Block diagram of the lock-in amplifier.

A very simplified schematic of a phase-sensitive detector is shown in Figure 4.18. This is generally referred to as a lock-in amplifier. The amplifier is of unity gain and is inverting. The trigger provides a voltage that is sufficient to switch the FETs. We see from the FET circuit that the output is connected alternately to the output of the amplifier or to the signal itself at a frequency equal to that of the square-wave output of the trigger. Thus the output is switched from an in-phase to an out-of-phase input signal at the reference oscillator frequency. The phase shifter accounts for phase lags in the experiment; for the time being we will assume that it has no effect.

Let us look at the output for some different situations.

Case 1. *Sine-wave input at a frequency equal to the reference frequency and in phase with the reference frequency*. Figure 4.19 shows the relationship between the input, the reference, output at (*a*), and the output at (*b*). The output at (*b*) has been filtered through the *RC* circuit, the time constant of which is larger than the period of the reference signal. We see that the device acts as a full-wave rectifier because during the positive half of the input wave cycle the output is connected through Q1 to the input, and for the negative half of the input wave cycle the output is inverted by the op-amp.

Case 2. Sine wave input at frequency equal to the reference frequency but $\pi/2$ *out of phase*. This situation is illustrated in Figure 4.20. The filtered output at (*b*) is now zero because the average at (*a*) is zero.

A derivation of the amplitude of the output signal is as follows. The input signal is given by

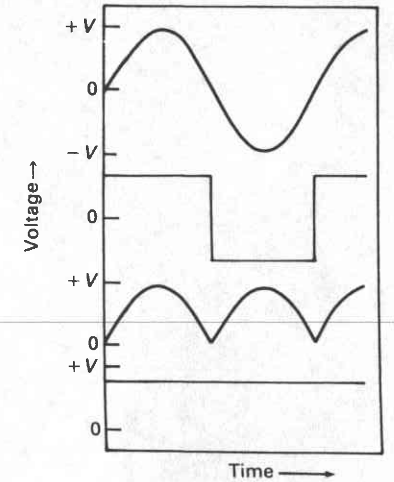
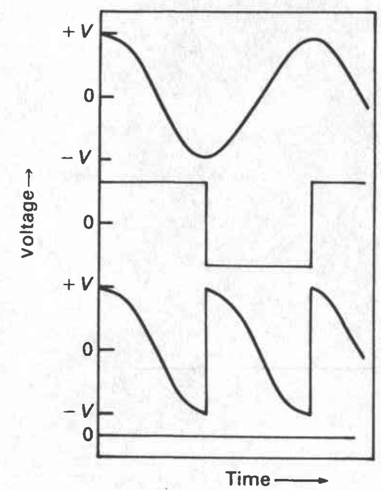


Figure 4.19. Case 1 as described in the text for the lock-in amplifier. The locations of points (*a*) and (*b*) are shown for the circuit in Figure 4.18.

where ψ is the phase difference between the input signal and the reference. We find that the output signal at (*b*) is

$$V_b = (2/\pi)V_0 \cos \psi \qquad (4.26)$$

where the $2/\pi$ factor results from the filtering process.



If the frequencies of the input and the reference are different (assumed, to begin with, by a small amount $\Delta\omega$), we write the input as

$$V_i = V_0 \sin(\omega + \Delta\omega)t \qquad (4.37)$$

This represents a slowly varying phase shift of the form

$$\psi = \Delta\omega t \qquad (4.28)$$

We therefore obtain a slowly varying output of the form

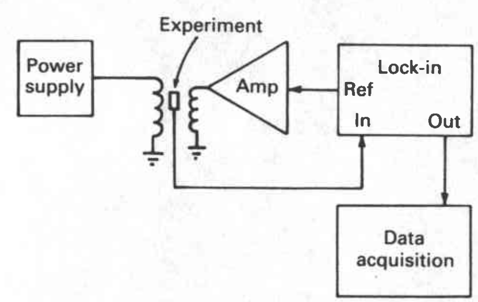
$$V_b = (2/\pi)V_0 \cos(\Delta\omega t) \qquad (4.29)$$

When the frequency difference is large, the oscillations in V_a are rapid and they are filtered out by the *RC* network. Hence the voltage at (*b*) is zero.

Thus the lock-in amplifier measures signals only at frequencies very close to the reference frequency, and it measures the in-phase components of those signals. Any noise present is not detected, because any component of the noise with the right frequency will have random phase.

Now all we need to do is to make sure that the signal we are detecting is at the right frequency and with the right phase. Phase problems due to wires, etc., can always be adjusted out with the phase shifter. Getting the signal at the right frequency is easy; we modulate the appropriate extrinsic parameter. Let us look at the example we gave for the MCA system—that of measuring some intrinsic quantity as a function of applied magnetic field. Consider the experimental system shown in Figure 4.21. A large magnetic field (the extrinsic quantity) is applied to the experiment. In addition a small oscillating magnetic field is applied by a small electromagnet (represented by the modulation coils). As a result of this modulation of the extrinsic parameter, the intrinsic parameter that is input into the lock-in amplifier is modulated at this same frequency as well. The reference signal from the lock-in is used to modulate the extrinsic quantity at the proper frequency. Thus the lock-in amplifier accepts only the signal due to the modulated intrinsic parameter and rejects all of the noise (which has either the wrong phase, or the wrong frequency, or both).

A convenient modification of the arrangement shown in Figure 4.21 is to



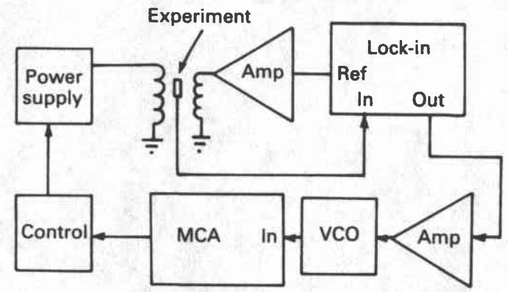


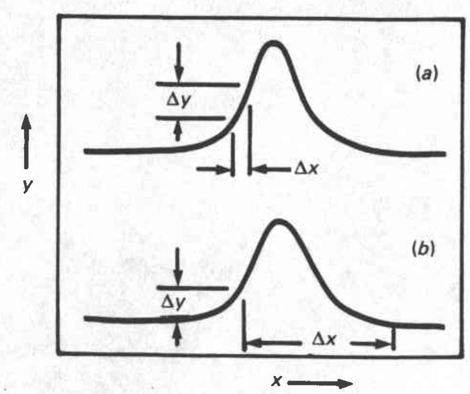
Figure 4.22. The use of signal averaging in conjunction with phase-sensitive detection.

use an MCA both to control the major component of the extrinsic quantity and to signal-average the data. Figure 4.22 shows how this could be done. Certainly the control of the extrinsic parameter may be modified in order to accommodate the purpose of the experiment.

So far we have implied that the modulation is small compared with the total range of extrinsic parameters of interest. This is not always the case. Two situations can arise: (1) the modulation is small (generally sinusoidal); (2) the modulation is a large square wave. We consider these two cases below.

a. *Small modulation*

Consider a hypothetical relationship between intrinsic and extrinsic parameters as shown in Figure 4.23. The range of x values chosen over which to make measurements is determined by the location of the interesting features in y, in this case the peak. Figure 4.23*a* shows that when the modulation



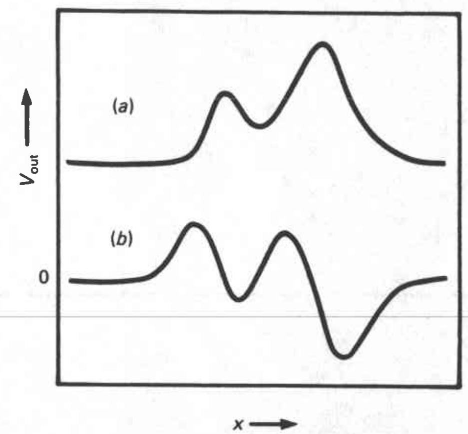


Figure 4.24. (*a*) The functional relationship $y = f(x)$ and (*b*) the signal measured by the lock-in amplifier using a small sinusoidal modulation.

signal is small the size of the change in the intrinsic quantity that is detected is related to the slope of the functional relationship between y and x. Thus we can write

$$\Delta y = \Delta x \frac{\partial y}{\partial x} \tag{4.30}$$

and the in-phase signal measured by the lock-in amplifier is

$$V_{out} = \frac{2}{\pi} \Delta x \frac{\partial y}{\partial x} \tag{4.31}$$

This means that we measure the first derivative of the signal. The lock-in output for the relationship shown in Figure 4.23*a* is illustrated in Figure 4.24.

b. *Large modulation*

As illustrated in Figure 4.23*b*, large modulation signals yield a Δy equal to $y(x)$. This is because a *zero* point on the $y = f(x)$ curve is used as a reference. Thus the in phase component of the lock-in amplifier output will look the same as the functional relationship between y and x. In this case we have to be sure that the shape of $y(x)$ is suitable and that Δx is large enough that we are referencing Δy to a zero point on the function.

From a practical standpoint it is important to make a proper choice of the modulation frequency. We cannot make the frequency too high in many cases because of the response time for the extrinsic parameter. If the best to avoid frequencies near 60 Hz or any of its low harmonics (i.e., 120 or 180 Hz). A few hundred hertz is generally suitable, but careful consideration should be made of the experimental details.

4.6 Digital-to-analog converters and analog-to-digital converters

Digital-to-analog (DAC) and analog-to-digital (ADC) converters comprise instruments that operate with digital signals to devices which operate with analog ones. While experiments generally deal with analog signals, it is nearly always necessary to convert these to digital signals if we want to process this information using a microcomputer, or for that matter even using an MCA. Most MCAs have ADCs built into their input. Conversely, if we want to use a computer to control an experiment it is possible in many cases to use digital instrumentation that connects directly to the digital output of the computer. Frequently, however, it is necessary to convert the digital output of the computer to an analog signal in order to, say, program a power supply. We begin with a description of the DAC, as this is the easiest to understand and will allow us to introduce some basic concepts of digital signal representation. As we shall see later, digital signals can be serial or parallel. For the time being, however, we will concentrate on parallel digital signals.

a. *The DAC (digital-to-analog converter)*

Any number may be represented in binary form. For the present discussion we consider the simple case of positive integers, but in general it should be understood that we are not necessarily limited to this case. The binary representation of a number is merely that number converted into base 2. For example the binary representation of the decimal number 37 would be 100101, that is,

$$1\times 32+0\times 16+0\times 8+1\times 4+0\times 2+1\times 1=37 \tag{4.32}$$

Although it takes many more digits to represent a number in binary form the digits need only be 0 or 1 rather than $0, 1, 2, \ldots, 9$ as in base 10. This makes things simple, since an electrical signal need only be either "high" (logical 1) or "low" (logical 0) to represent the digit. Each digit in the binary number is referred to as a bit. Hence 100101 is a six-bit binary number. It is easy to see that an n-bit binary number can represent integers from 0 up to 2^n-1. We can represent a digital quantity in the electrical sense by a series of n wires each with either a high or low voltage (logical 1 or 0). There is, of course, a common ground connection.

Figure 4.25 shows a simple circuit for converting this kind of digital signal into an analog one: an 8-bit DAC is illustrated. Although 8 is not always the number of bits, it is not uncommon. The larger the number of bits the

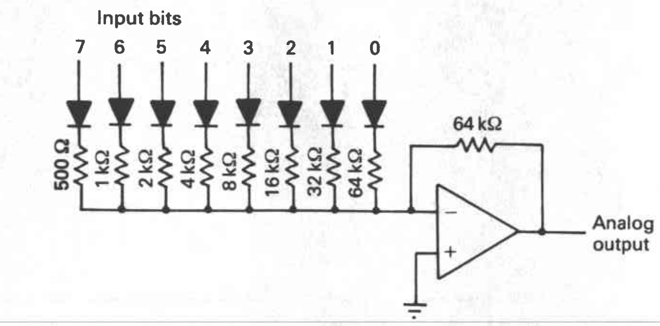


Figure 4.25. The digital-to-analog converter (DAC).

that the analog output signal is given in terms of the signals b_i on the eight input bits ($i=0$ to 7), where $b_i=0$ or 1 is represented by 0 V or $+V$, respectively. The output is found to be

$$V_{out}=-V\sum_{i=0}^{n-1}2^i b_i \tag{4.33}$$

We see that V_{out} cannot change continuously but is "quantized" in steps of V; the smallest change in V_{out} occurs for a change in b_0 from 0 to 1 (or from 1 to 0).

The maximum output occurs for all $b_i=1$ and this means that the resolution will be 1 part in 2^n-1. The larger the value of n, of course, the better the resolution. A circuit of the type shown in Figure 4.25 provides a convenient method of converting a parallel digital signal (see Section 4.7) from a microcomputer to an analog signal for experimental control. The number of bits of the DAC that you use should be consistent with the number of parallel bits that can be supplied as output by your computer. We shall say more on this in the section on computer interfacing. DACs can be bought for a few dollars, so if you are interested in having something that works well, it is best to buy one. If you are interested in studying how the DAC works, it is useful to build the circuit in Figure 4.25. The F/V (frequency to voltage converter) discussed in Section 3.2 is sometimes called a *time-domain DAC.*

b. *The ADC (analog-to-digital converter)*

There are a variety of methods of converting analog signals to digital signals; these are in general more complex than the methods for converting digital to analog. We describe one simple method.

Figure 4.26 shows the block diagram of a simple ADC. This is a 4-bit ADC, although in practice ADCs are typically of 8 or more bits. A 4-bit

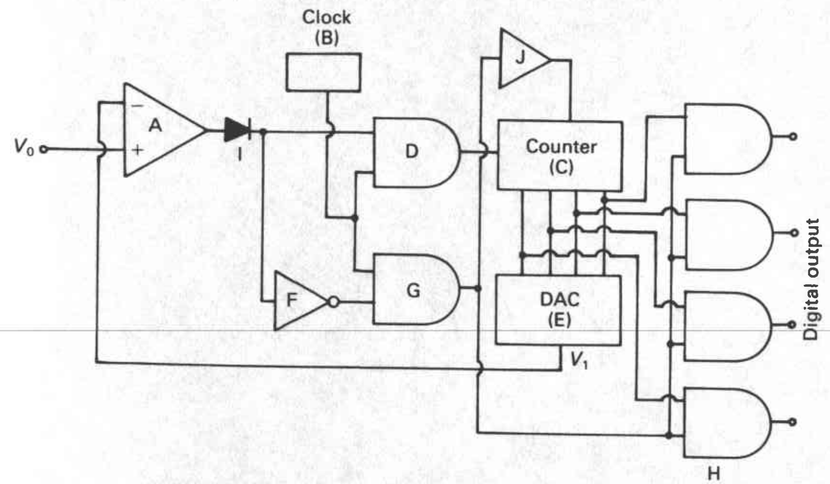


Figure 4.26. A 4-bit analog-to-digital converter (ADC): A = comparator (differential amplifier); B = clock; C = binary (4-bit) counter; D = clock gate; E = digital-to-analog converter (4-bit); F = inverter, G = reset gate; H = output gates; I = diode; J = buffer.

works. A clock B produces pulses at some fixed frequency. These are counted by a 4-bit binary counter C, beginning at 0000, and progressing 0001, 0010, 0011 etc. The digital output of the counter is converted into an analog signal by a DAC. This analog signal V_1 is input into a comparator A (differential amplifier). The other input of the comparator is the analog signal itself (the one we want to digitize)—a voltage V_0. As long as $V_o > V_1$ the comparator outputs a positive voltage and this is allowed to pass the diode I. This signal means that gate D is open whenever a clock pulse is present. Thus as long as $V_0 > V_1$ then the binary counter continues to count; as the counter counts, V_1 increases in steps. The size of the step is determined by the details of the DAC (mostly by the number of bits). This continues until $V_1 \geq V_0$ at which time the output of the comparator is 0 or negative, and this is blocked by the diode. Hence gate D is turned off and will no longer pass clock pulses, so that the counter stops counting. This logical "0" from the comparator is inverted by the inverter F to a logical "1". This holds gate G open for one clock pulse. This clock pulse does two things; (1) it opens gates H on which are the logical signals from the counter and outputs the digital (binary) signal; (2) it resets the counter to 0000. Buffer J ensures that the digital signal is output before the counter is reset. Thus the digital signal that is output is the one for which its analog counterpart (as produced by the DAC) is equal to the analog input signal (as determined by the comparator). Hence the output is the digital

The speed at which the ADC samples the analog input and produces a digital output depends on the speed of the clock and the number of bits in the counter. The number of bits is chosen according to the resolution required. The speed of the clock must be chosen to be compatible with the speed of the other electronics in the circuit.

The successive-approximation ADC (a faster kind of ADC). In this method a similar comparison method is used, but the binary counter is replaced by a programmable chip. This chip first sets the MSB = 1 and all other bits to zero. It then compares an analog signal V_1 to V_0. If $V_0 < V_1$ the MSB is set to zero and the next most significant bit is set to 1; otherwise the MSB is left equal to 1 and the next most significant bit is set to 1 as well. Each bit is "inspected" in this way; it is rather like a binomial search process. This process requires *n* inspections for *n* bits, while on average the previous method requires

$$N = 2^{n-1} \tag{4.34}$$

inspections. Thus the successive-approximation ADC is faster by a factor of $2^{n-1}/n$. For $n = 8$ this is a factor of 16, while for $n = 12$ it is a factor of 170, and for $n = 16$ the factor is over 2000.

Typically, we would choose a complete ADC in the form of an integrated circuit. Consideration must be given to what the digital signal will be used for and we must ensure that the output is compatible with the other devices in terms of voltage, speed, and the use of "handshaking" lines. We discuss these aspects further in the following section.

4.7 Standard digital-to-digital interfaces

In Section 4.6 we discussed the conversion of analog signals to digital signals and vice versa. These techniques are useful for converting analog signals from an experiment to digital form for input into a computer and for converting digital signals from a computer to analog form for instrument control. In this section we discuss the connection of equipment with specific digital interfaces to a microcomputer. These digital interfaces are either serial or parallel in nature. Serial signals are those in which the bits are transmitted one after the other along a single electrical connection. Parallel signals are those in which *n* bits are transmitted simultaneously along *n* electrical connections. As we shall see, each of these systems has its own advantages and disadvantages. In the past there has been a good deal of confusion over interfacing standards because of the wide variety of techniques used by different manufacturers. It is certainly not possible within the scope of this book to discuss all of the interfacing methods used in recent years. Fortunately, the computer industry now seems to be agreeing upon a fairly small number of interfacing standards. Two of these, for connecting microcomputers to laboratory equipment. We consider these two in some detail as they are both useful and illustrate nicely the differences between serial and parallel interfacing. However, if you should encounter a piece of equipment or a microcomputer that has a different serial or parallel interface (a situation that is not unlikely), a knowledge of how RS-232 and IEEE-488 work should make it easy for you to decipher the necessary instruction manuals and understand how to use the interface.

a. RS-232 serial interface

Signals sent in the RS-232 standard represent characters (numbers, etc.) by a series of logical bits (1s and 0s). The voltage levels used on an RS-232 line to represent logical 1 and 0 are shown in Figure 4.27. In most cases characters sent on RS232 lines are represented by the ASCII code (*A*merican *S*tandard *C*ode for *I*nformation *I*nterchange). This code represents each character by a *seven*-bit binary number. In transferring data we are primarily concerned with the transmission of numbers but it is obvious that other characters (spaces, decimal points, line feeds, returns, etc.) are necessary as well. Table 4.2 gives the full ASCII code.

We see that the ten digits 0 to 9 correspond to the binary representation 0110000 to 0111001. These bits are transmitted at a particular frequency. This is known as the baud rate and specifies the number of bits transmitted per second. Standard baud rates for data transmission are 75, 110, 134.5, 150, 300, 600, 1200, 2400, 4800, 9600 and 19200. As a character consists of the 7-bit ASCII code and is generally has appended by one or two

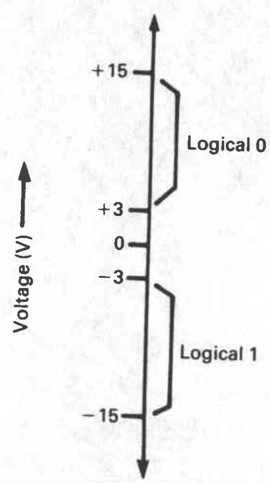


Table 4.2. ASCII character set, in which the characters are represented by seven bits $b_6b_5b_4b_3b_2b_1b_0$: b_0 is the 2^0 bit, b_1 is the 2^1 bit, etc.

	$b_6b_5b_4$							
$b_3b_2b_1b_0$	000	001	010	011	100	101	110	111
0000	NUL	DLE	SP	0	@	P	`	p
0001	SOH	DC1	!	1	A	Q	a	q
0010	STX	DC2	"	2	B	R	b	r
0011	ETX	DC3	#	3	C	S	c	s
0100	EOT	DC4	$	4	D	T	d	t
0101	ENQ	NAK	%	5	E	U	e	u
0110	ACK	SYN	&	6	F	V	f	v
0111	BEL	ETB	'	7	G	W	g	w
1000	BS	CAN	(	8	H	X	h	x
1001	HT	EM	)	9	I	Y	i	y
1010	LF	SUB	*	:	J	Z	j	z
1011	VT	ESC	+	;	K	[	k	{
1100	FF	FS	,	<	L	\	l	|
1101	CR	GS	–	=	M	]	m	}
1110	SO	RS	.	>	N	^	n	~
1111	SI	US	/	?	O	—	o	DEL

Key: Mnemonics and functions of the control codes (ASCII 0000000 to 0011111)

NUL	Null	FF	Form feed	ETB	End transmission block
SOH	Start of heading	CR	Carriage return	CAN	Cancel
STX	Start of text	SO	Shift out	EM	End of medium
ETX	End of text	SI	Shift in	SUB	Substitute
EOT	End of transmission	SP	Space	ESC	Escape
ENQ	Enquiry	DLE	Data link escape	FS	File separator
ACK	Acknowledge	DC1	Device control 1	GS	Group separator
BEL	Bell	DC2	Device control 2	RS	Record separator
BS	Backspace	DC3	Device control 3	US	Unit separator
HT	Horizontal tab	DC4	Device control 4	DEL	Delete
LF	Line feed	NAK	Negative acknowledge		
VT	Vertical tab	SYN	Synchronize		

information bits on either end (we discuss these shortly), the maximum rate of character transmission is about 1/10th of the baud rate.

Whenever data are not being transmitted, the RS-232 line is in the logical 1 state; this is the idle state. When data are transmitted each 7-bit character is proceeded by a start bit. This is *one* logical 0 bit. This is followed by the seven character bits. After this comes an optional parity bit. Finally there are one or two stop bits. The stop bits are logical 1 bits. It is important to note that when a character is transmitted, the lowest-order bit, b_0 (see Table 4.2), is transmitted first.

The parity may be either even or odd and is determined by the number of logical 1 bits in the transmitted character. Table 4.3 shows some examples of the parity bit. Some examples of ASCII data transmission are illustrated